The impact of ignoring multiple membership data structures in multilevel models

Hyewon Chung1* and S. Natasha Beretvas2

1Department of Psychology, John Jay College of Criminal Justice, City University of New York, USA
2Department of Educational Psychology, University of Texas at Austin, USA

This study compared the use of the conventional multilevel model (MM) with that of the multiple membership multilevel model (MMMM) for handling multiple membership data structures. Multiple membership data structures are commonly encountered in longitudinal educational data sets in which, for example, mobile students are members of more than one higher-level unit (e.g., school). While the conventional MM requires the user either to delete mobile students’ data or to ignore prior schools attended, MMMM permits inclusion of mobile students’ data and models the effect of all schools attended on student outcomes. The simulation study identified underestimation of the school-level predictor coefficient, as well as underestimation of the level-two variance component with corresponding overestimation of the level-one variance when multiple membership data structures were ignored. Results are discussed along with limitations and ideas for future MMMM methodological research as well as implications for applied researchers.

1. Introduction

Multilevel models (MMs), also known in the literature as hierarchical linear models, mixed effects models, or random coefficient regression models (Goldstein, 2010; Raudenbush & Bryk, 2002; Snijders & Bosker, 2011), are commonly used to handle clustered data structures. Examples of clustered data are scenarios in which students are nested in schools, or patients within hospitals. The traditional MM can be used to handle pure hierarchical data structures, in which each lower-level unit belongs to a single higher-level unit. Figure 1 shows a network graph, also referred to as a unit diagram (Browne, Goldstein, & Rasbash, 2001; Goldstein, 2010; Rasbash & Browne, 2001), which illustrates the pure clustering of students within high schools where the lines connecting level-one with level-two units do not cross. In practice, however, it is unrealistic to assume that all multilevel data are purely hierarchical.

*Correspondence should be addressed to Hyewon Chung, Department of Psychology, John Jay College of Criminal Justice, 445 West 59th Street, New York, NY 10019, USA (e-mail: hchung@jjay.cuny.edu).

DOI:10.1111/j.2044-8317.2011.02023.x
A typical example of a non-pure hierarchical data structure is frequently encountered when analysing longitudinal data sets. For example, the National Educational Longitudinal Survey (NELS: 88) data set tracked participants in 8th, 10th and 12th grades. An analysis of the NELS: 88 data set revealed that approximately 10% of the sampled students reported changing schools at least once (i.e., attended multiple schools). Thus, the data set has a multiple membership structure (Fielding & Goldstein, 2006; Fielding, Thomas, Browne, Leyland, Spencer, & Davison, 2006; Hill & Goldstein, 1998; Rasbash & Browne, 2008). Figure 2 depicts a multiple membership data structure in which some students are affiliated with more than one high school. The lines in the network graph in Figure 2 cross, indicating a non-pure hierarchy. In addition, more than one line connects some students with a school, indicating a multiple membership structure. Other sources of multiple membership structures are encountered in different social and behavioural science research including medical (Browne et al., 2001; Chandola, Clarke, Wiggins, & Bartley, 2005), socio-economic (Goldstein, Rasbash, Browne, Woodhouse, & Poulain, 2000), and educational research (Browne et al., 2001; Goldstein, Burgess, & McConnell, 2007; Leckie, 2009). The current study, however, will employ school mobility as its source of a multiple membership data structure.

Use of the MM involves the assumption that lower-level units are purely nested in higher-level units. The multiple membership multilevel model (MMMM) is designed specifically to handle the complexity of multiple membership data structures. Most applied researchers, however, employ ad hoc procedures that remove or ignore the problematic multiple membership data structures. In educational research, for example, Lee (2000) assessed how high school size affected students’ academic development using a conventional MM. The author restricted samples to students who attended the same high schools between 10th and 12th grade, thereby limiting the generalizability of the findings to the population of students who attended the same school over the time period of the study.

Alternatively, when handling mobility, some researchers have included student mobility as a student-level predictor, indicating whether a student is mobile or non-mobile, and yet ignored the set of schools attended and recognized only one (typically the most recent) of the schools that each mobile student attended (South, Haynie, & Bose, 2007). As with the deletion strategy, this approach fails to take into account the characteristics of the previous school(s) that the mobile students attended and
their contribution to mobile students’ academic achievement during the period of analysis. This threatens the validity of inferences about schools’ effects on student-level outcomes.

Although most applied studies have used one of these strategies to handle student mobility, a few studies have employed the MMMM. For example, Leckie (2009) conducted a study to investigate the effect of student mobility on academic achievement and compared models that ignored the multiple membership structure with models that included it. The findings revealed that the model ignoring the multiple membership within secondary schools and neighbourhoods resulted in smaller school- and neighbourhood-level variance component estimates. In addition to comparing parameter estimates, Leckie used the models’ deviance information criterion (DIC; Spiegelhalter, Best, Carlin, & van der Linde, 2002) values to compare the relative fit of each model.

The DIC is used in Bayesian model selection to calculate the relative fit of each model (Spiegelhalter et al., 2002). The larger the DIC is, the worse the fit. The rules of thumb for using the DIC to compare models’ fit are based on recommendations for the AIC index presented in Burnham and Anderson’s (1998) work and cited in Spiegelhalter et al., (2002, p. 613): ‘models receiving AIC within 1–2 of the “best” deserve consideration, and 3–7 have considerably less support’. Researchers have suggested minimum differences of either 5 (Li, Bolt, & Fu, 2006) or 10 (Leckie, 2009) as cut-offs for substantial decreases in model fit. With his real data set, Leckie (2009) had found that the MMMM consistently resulted in lower DIC values, supporting its preferred fit over the MM.

Multiple membership data structures are commonly encountered and yet inappropriately handled in educational research. Therefore, the effect of these multiple membership data structures on parameter estimation should be empirically assessed and more explicitly demonstrated so that applied researchers can better understand the need for these models. While prior research has compared results between appropriately modelling and ignoring multiple membership data structures, these studies have involved real data sets. When comparing parameter estimates using real data sets, one cannot assess which method produces the more accurate estimates of the ‘true’ parameters as these true parameters are not known. Very few studies were found that involved a simulation study designed to assess MMMM parameter recovery. Browne et al. (2001) conducted a simulation, but they focused solely on assessing recovery of the true random effects variance component. The current study is designed to build upon the simulation research of Browne et al. and also included an empirical assessment of the MMMM’s parameter recovery of level-one and level-two predictors’ coefficients. Additionally, the current study compares parameter recovery when the MMMM is used with multiple membership data structures against when MM is used. Thus, this research is intended to demonstrate the impact of ignoring multiple membership structures and to assess how well the MMMM’s parameter values are recovered under a variety of conditions.

2. Method

A simulation study was designed to compare two methods used for handling multiple membership data structures. Use of the MMMM that appropriately modelled the multiple membership data structure was compared with use of the MM that ignored the multiple membership data structure. The latter model recognized only the most recent level-two clustering unit (here, school) of which the level-one unit (here, student) was a member.
2.1. Conditions

The conditions manipulated in the simulation study were: the percentage of students who changed schools (10%, 20%), the intra-class correlation (ICC; 5%, 15%, 25%), the number of schools (50, 100), the number of students per school (30, 60), and the number of schools attended by mobile students (2, 3).

Several large-scale national longitudinal data sets including primary school students and secondary school students (NELS: 88, NELS: 2000, and ECLS-K) were examined in order to select reasonable values for degrees of student mobility. The results indicated a range of student mobility from approximately 8% to 17% over three school years. To mimic findings from the real data sets, values of 10% and 20% were selected to depict small and moderate levels of student mobility.

The ICC value here refers to a conditional ICC – the correlation between two level-one units within a level-two unit after controlling for level-one and level-two covariates. The conditional ICC values from the analyses conducted using two longitudinal data sets (NELS: 88 and ECLS-K) ranged from 11.1% to 23.3%. Conditional ICC values used in previous methodological research on multilevel models have typically ranged from 5% to 30%. Three conditional ICC values were thus used (5%, 15%, and 25%) to match the values found in previous methodological and applied research.

The sample sizes used in the current simulation study were chosen to compare the results from previous related methodological studies and to assess a minimum sample size necessary for parameter recovery in MMMM scenarios. The optimal minimum level-two sample size has been debated by several methodological researchers (Kreft & de Leeuw, 1998; Maas & Hox, 2005; Van der Leeden, Busing, & Meijer, 1997). To assess the impact of the level-two sample size in MMMM parameter recovery, two values for the number of schools were investigated. Maas and Hox noted that educational data sets commonly include around 50 schools. Van der Leeden et al. (1997) found that a minimum of 100 level-two units was necessary for reasonable parameter recovery. The current study thus assessed how well parameters are recovered for data sets containing 50 and 100 schools.

Two values for the number of level-one units per level-two unit sample size (here, corresponding to the average school size) were selected based on values used in previous methodological studies on multilevel models. In such research, the typical average number of students sampled per school ranges from 5 to 61 (e.g., Browne & Draper, 2006; Maas & Hox, 2005). Maas and Hox suggested that a school size of 30 was typical in educational research. The current study is intended to focus on the impact of mobility on model estimation under conditions with sufficient sample sizes and thus values lower than 30 were not investigated. Values of 30 and 60 for the average school size were investigated.

2.2. Data generation

In the current study, the notation for the MMMM will match that used by Beretvas (2010) which is based on the levels formulation of Raudenbush and Bryk (2002). The corresponding formulation of the MMMM of interest using the notation of Browne et al. (2001) appears in the Appendix.

The data were generated to fit a two-level multiple membership data structure to mimic the clustering of students at level one within schools at level two, with some students as members of multiple schools (mm% of the students). MLwiN software (version 2.15, Rasbash, Charlton, Browne, Healy, & Cameron, 2009) was used to generate
1,000 data sets per combination of conditions. The data were generated to include one student and one school predictor using the level-one equation

\[ Y_{i(j)} = \beta_0(j) + \beta_1(j)X_{i(j)} + e_{i(j)}, \]  

(1)

where \( Y_{i(j)} \) represents the outcome for student \( i \) who attended the set of schools \( \{j\} \), and \( e_{i(j)} \sim N(0, \sigma^2) \) represents the deviation of student \( i \)'s score from the mean outcome of \( \{j\} \) after adjusting for the student predictor, \( X_{i(j)} \). The level-two equation was

\[
\begin{align*}
\beta_0(j) &= \gamma_{00} + \gamma_{01} \sum_{b \in \{j\}} w_{ib} Z_b + \sum_{b \in \{j\}} w_{ib} u_{0b}, \\
\beta_1(j) &= \gamma_{10},
\end{align*}
\]

(2)

where \( u_{0b} \sim N(0, \tau_{00}) \) is the residual for unit \( b \in \{j\} \), after controlling for the weighted average of the set of \( \{j\} \) schools’ values on \( Z_b \) and the student predictor, \( X_{i(j)} \), and \( w_{ib} \) represents student \( i \)'s weight associated with unit \( b \in \{j\} \). The weights for each unit \( i \) associated with the set of level-two units, \( \{j\} \), in equation (2) sum to one, that is, \( \sum_{b \in \{j\}} w_{ib} = 1 \). The coefficient, \( \gamma_{10} \), for the student predictor, \( X \), was modelled as fixed across level-two units in equation (2).

There are several ways to assign weights in a multiple membership scenario. One method involves assigning equal weights to each school attended regardless of the proportion of time that a student attended the school. Another method involves assigning weights that reflect the length of time a student attended each school, potentially resulting in unequal weights. Alternatively, weights could be assigned using a combination of factors that might contribute to a school’s hypothesized effect on a student outcome. For example, a researcher might hypothesize that the length of time that a student attends a school and how recently that school was attended should both contribute to the weight (resulting in a more recently attended school being assigned more weight). Ultimately, a researcher might be unsure of the optimal weights to use for an MMM and could use a variety of weight assignment strategies and compare the fit of the resulting estimated models.

In the current study the second strategy was employed, simulating a scenario in which school enrolment was available at three different time points. If a student was a member of only one school across the three time periods, then a weight of 1 was associated with the relevant school. For students attending three different schools across the three time points, each school was assigned a weight of 1/3. In conditions where the number of schools attended by the ‘mobile’ student was two, unequal weights were assigned such that one school was assigned a weight of 1/3 and the other school was assigned a weight of 2/3.

The generating values for the fixed effects coefficients were 100 for the intercept (\( \gamma_{00} \)), 0.4 for \( X \) (\( \gamma_{10} \)), and 0.4 for \( Z \) (\( \gamma_{01} \)). Values for the level-one and level-two predictors (i.e., for \( X \) and \( Z \), respectively) were sampled from normal distributions with means of 50 and standard deviations of 10. These values were selected based on values used in previous multilevel model simulation studies (Maas & Hox, 2005; Meyers & Beretvas, 2006). The level-one and level-two residuals (\( e_{i(j)} \) and \( u_{0(j)} \), respectively) were also sampled from normal distributions with means of 0. The variance of the \( e_{i(j)} \) was generated to be 1 across conditions. The generating values for \( \tau_{00} \) (the variance of the level-two residuals, \( u_{0(j)} \)) were 0.05265, 0.17647, and 0.33333 for conditions with conditional ICC values of 5%, 15% and 25%, respectively. Once the values for the level-one
and level-two residuals and predictors were generated for each level-one and level-two unit, the values were substituted into the relevant equation, (1) or (2), to obtain the value on the outcome variable, \( Y \), for each simulated student.

2.2.1. Generating multiple membership patterns

The data structure was generated to mimic the multiple membership data structure that results from students changing schools. Two maximum values \( (m = 2, 3) \) for the number of level-two units, of which each level-one unit was a member, were investigated. This meant that in one set of conditions \( (m = 2) \), mobile students attended two different schools across the three time points. In the other set of conditions \( (m = 3) \), mobile students attended a different school at each of the three time points.

These values for \( m \) were fully crossed with the percentage of mobile students. The two values for the percentage of mobile students that were investigated here were \( mm\% = 10\% \) and \( 20\% \). Thus, the condition's percentage of multiple members \( (mm\%) \) was used to identify which level-one units were associated with \( m \) level-two units. And \( 100 - mm\% \) of the level-one units (i.e., non-mobile students) were associated with only a single level-two unit (school). The non-mobile students were randomly assigned to one school. Mobile students \( (mm\%) \) were randomly assigned to \( m \) of the 50 (or 100, depending on the condition) different schools. The MLwiN relevant macro can be provided upon request by the first author.

2.3. Analyses

To be consistent with previous studies employing Markov chain Monte Carlo (MCMC) in estimating MMMMs, MCMC (Browne, 2009; Rasbash, Steele, Browne, & Goldstein, 2009), as implemented in the MLwiN software (including use of the default priors), was used to estimate each model. One chain was run with 50,000 iterations with a burn-in of 5,000. In a pilot study, estimation of each fixed and random effects variance estimate was checked using the Raftery–Lewis statistic (Raftery & Lewis, 1992).

Two different models were fitted to each multiple membership data set that was generated. The MM that ignores multiple membership was estimated and the correct MMMM was also estimated. The two models’ results were compared to evaluate the impact of ignoring and of appropriately modelling multiple membership data structures.

When estimating the MM that ignored the multiple membership data structure, only the effect of the last school attended was modelled. Thus students (level one) were considered purely nested within the most recent school attended (level two). The level-one equation for the mismatched multilevel model was

\[
Y_{ij} = \beta_{0j} + \beta_{1j} X_{ij} + e_{ij},
\]

and the level-two equation was

\[
\begin{cases}
\beta_{0j} = \gamma_{00} + \gamma_{01} Z_j + u_{0j}, \\
\beta_{1j} = \gamma_{10},
\end{cases}
\]

where \( Y_{ij} \) represents the outcome for student \( i \) from school \( j \) at the third time point. Note that the student-level predictor \( (X) \) and school-level predictor \( (Z_j) \) at the third time
Ignoring multiple membership

point were included in the model. Note also that the effect of the student characteristic (X) was modelled as fixed across schools (see (4)), matching the assumption in the generating (and estimating) MMMM.

Parameter estimates and DIC values were saved for each of the 1,000 replications for each combination of conditions under each of the MM and MMMM. The proportion of replications was tallied in which the difference between the MMMM and MM DIC values was less than 5. SAS software was used to summarize the estimated fixed and random effects parameters across the 1,000 simulated data sets per condition. The parameters’ recovery was evaluated using the relative parameter bias given by

\[ B(\hat{\theta}_i) = \frac{\hat{\theta}_i - \theta_i}{\theta_i}, \]

where \( \theta_i \) is the generating true value of \( i \)th parameter, and \( \hat{\theta} \) is the average of the estimates for the \( i \)th parameter across the 1,000 simulated data sets per condition. Hoogland and Boomsma (1998) suggested cut-offs for minimum relative parameter bias values (of magnitude 5%) that can be considered indicators of substantial bias.

Finally, analyses of variance (ANOVAs) were used to explore which simulation condition factors affected the relative bias of parameter estimates. Given the very large data set (including 1,000 replications for each of the 48 conditions), statistical significance would not necessarily indicate that a factor had a practically significant effect. Thus, only factors that were statistically significant (\( p < .01 \)) predictors of relative bias and that were associated with effect size measure values (\( \eta^2 \)) greater than .01 were used to identify which condition factors had practically significant effects. Other multilevel model simulation studies (Krull & MacKinnon, 1999) have used a similar procedure for identifying significant factors affecting parameter bias.

3. Results

3.1. Fixed effects estimates

No substantial relative parameter estimation bias was detected for either the intercept estimates or the level-one predictor coefficient. Means and standard deviations of the relative bias of the intercept across all 48 conditions were \( M = 0.018 \) and \( SD = 0.007 \) for MM estimates, and \( M = -0.00003 \) and \( SD = 0.003 \) for MMMM estimates. Means and standard deviations of the relative bias of the level-one predictor coefficient were \( M = 0.00002 \) and \( SD = 0.007 \) for MM estimates, and \( M = -0.00003 \) and \( SD = 0.005 \) for MMMM estimates. Although no substantial bias was found under the MM and MMMM, the relative bias for the intercept was less for the MMMM.

When the MMMM was estimated, no substantial bias was found in the level-two predictor coefficient estimate across all conditions. The mean and standard deviation of the relative bias were \( M = 0.00002 \) and \( SD = 0.015 \) for the MMMM intercept estimates. However, failure to model the effects of mobile students’ prior schools led to underestimation of the level-two predictor’s fixed effect (\( M = -0.088 \) and \( SD = 0.036 \)).

Given the identification of substantial bias in the MM estimates, an ANOVA was conducted to explore the factors that affected the relative parameter bias of the MM estimates. The ANOVA results revealed that the percentage of mobile students (see Tables 1 and 2), \( mm\% \), was very strongly associated with the observed negative relative bias, \( F(1, 47,980) = 182,474.00, p < .001, \eta^2 = .684 \). The average relative bias was \( -0.118 \) in conditions with 20% mobile students, compared to \( -0.059 \) in conditions with 10%
Table 1. Summary of mean relative parameter bias by condition for MM parameter estimates of the level-two predictor, $\gamma_{01}$, level-one variance component, $\sigma^2$, and level-two variance component, $\tau_{00}$

<table>
<thead>
<tr>
<th>Condition</th>
<th>Parameter</th>
<th>$\gamma_{01}$</th>
<th>$\sigma^2$</th>
<th>$\tau_{00}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICC</td>
<td>0.05</td>
<td>$-0.088$</td>
<td>1.332</td>
<td>$-0.197$</td>
</tr>
<tr>
<td></td>
<td>0.15</td>
<td>$-0.088$</td>
<td>1.340</td>
<td>$-0.151$</td>
</tr>
<tr>
<td></td>
<td>0.25</td>
<td>$-0.088$</td>
<td>1.353</td>
<td>$-0.145$</td>
</tr>
<tr>
<td>$c$</td>
<td>50</td>
<td>$-0.089$</td>
<td>1.338</td>
<td>$-0.163$</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>$-0.088$</td>
<td>1.346</td>
<td>$-0.165$</td>
</tr>
<tr>
<td>$n$</td>
<td>30</td>
<td>$-0.088$</td>
<td>1.343</td>
<td>$-0.177$</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>$-0.089$</td>
<td>1.340</td>
<td>$-0.151$</td>
</tr>
<tr>
<td>$m$</td>
<td>2</td>
<td>$-0.076$</td>
<td>1.247</td>
<td>$-0.140$</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>$-0.101$</td>
<td>1.436</td>
<td>$-0.189$</td>
</tr>
<tr>
<td>$mm%$</td>
<td>10%</td>
<td>$-0.059$</td>
<td>0.933</td>
<td>$-0.103$</td>
</tr>
<tr>
<td></td>
<td>20%</td>
<td>$-0.118$</td>
<td>1.750</td>
<td>$-0.226$</td>
</tr>
</tbody>
</table>

*Note.* ICC = intra-class correlation, $c$ = number of schools, $n$ = school size, $m$ = number of schools that mobile students attended, and $mm\%$ = percentage of mobile students.

Mobile students. This indicates that the more mobile students in the data set, the stronger the negative relative bias for the MM level-two predictor coefficient estimates.

The number of schools that mobile students attended, $m$, was also found to have a substantial effect on the observed negative relative bias, $F(1, 47,980) = 32,477.60$, $p < .001$, $\eta^2 = .122$. The average relative bias in the conditions where mobile students attended two different schools across the three time measurement points was $-0.076$. The average relative bias in the conditions where mobile student attended three different schools was $-0.101$. The more schools that mobile students attended, the stronger the negative relative bias.

In addition to the main effect, the interaction between the degree of mobility and the number of schools attended by mobile students was found to be substantially related to the observed negative bias, $F(1, 47,980) = 3,670.60$, $p < .001$, $\eta^2 = .014$. In the condition with 10% mobility and mobile students attending two schools, the average bias was $-0.051$, as compared to $-0.067$ when the 10% of mobile students attended three schools. In the conditions with 20% mobility, the average bias estimates were $-0.101$ and $-0.135$ for mobile students attending two and three schools, respectively. In other words, as the number of schools attended increased, the difference in bias that was a function of percent mobility tended to become slightly larger. No other main effect or two-way interactions were found to have substantial effects on the relative parameter bias.

3.2. Random effects variance component estimates

MMMM estimates of the level-one variance component were not substantially biased ($M = 0.0013$, $SD = 0.0278$). Under the MM, however, the results revealed that inappropriate modelling of multiple membership data structures tended to result in significant overestimation of the level-one variance component ($M = 1.3415$, $SD = 0.4971$). An ANOVA revealed that two main effects factors were found to be
Table 2. Analysis of variance results for the relative bias of the MM estimates of the level-two predictor, $\gamma_{01}$, Level-one variance component, $\sigma^2$, and level-two variance component, $\tau_{00}$

<table>
<thead>
<tr>
<th>Source</th>
<th>$\gamma_{01}$</th>
<th>$\sigma^2$</th>
<th>$\tau_{00}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SS F-ratio $\eta^2$</td>
<td>SS F-ratio $\eta^2$</td>
<td>SS F-ratio $\eta^2$</td>
</tr>
<tr>
<td>Mobility</td>
<td>41.837 182,474.00*</td>
<td>8,020.884 113,807.00*</td>
<td>181.145 2,555.15*</td>
</tr>
<tr>
<td>ICC</td>
<td>&lt;.001 0.04 &lt;.001</td>
<td>3.694 26.21* &lt;.001</td>
<td>26.191 184.72* .007</td>
</tr>
<tr>
<td>Number of schools</td>
<td>0.005 22.80* &lt;.001</td>
<td>0.766 10.87* &lt;.001</td>
<td>0.047 0.66 &lt;.001</td>
</tr>
<tr>
<td>School size</td>
<td>&lt;.001 2.05 &lt;.001</td>
<td>0.092 1.31 &lt;.001</td>
<td>8.181 115.40* .002</td>
</tr>
<tr>
<td>$m$</td>
<td>7.446 32,477.60*</td>
<td>430.220 6,104.33*</td>
<td>28.838 406.77* .008</td>
</tr>
<tr>
<td>Mobility $\times$ ICC</td>
<td>&lt;.001 &lt;.001 &lt;.001</td>
<td>0.410 2.91 &lt;.001</td>
<td>1.233 8.70* &lt;.001</td>
</tr>
<tr>
<td>Mobility $\times$ Number of schools</td>
<td>0.001 5.51 &lt;.001</td>
<td>0.073 1.03 &lt;.001</td>
<td>0.121 1.70 &lt;.001</td>
</tr>
<tr>
<td>Mobility $\times$ $m$</td>
<td>0.842 3,670.60*</td>
<td>21.153 300.14*</td>
<td>2.307 32.54* .001</td>
</tr>
<tr>
<td>ICC $\times$ Number of schools</td>
<td>&lt;.001 0.09 &lt;.001</td>
<td>0.003 0.02 &lt;.001</td>
<td>0.473 3.33 &lt;.01</td>
</tr>
<tr>
<td>ICC $\times$ School size</td>
<td>&lt;.001 0.03 &lt;.001</td>
<td>0.001 50.32 &lt;.001</td>
<td>7.135 50.32* .002</td>
</tr>
<tr>
<td>ICC $\times$ $m$</td>
<td>&lt;.001 0.15 &lt;.001</td>
<td>0.055 0.03 &lt;.001</td>
<td>0.004 0.03 &lt;.01</td>
</tr>
<tr>
<td>Number of schools $\times$ School size</td>
<td>&lt;.001 0.02 &lt;.001</td>
<td>0.001 0.63 &lt;.001</td>
<td>0.045 0.63 &lt;.001</td>
</tr>
<tr>
<td>Number of schools $\times$ $m$</td>
<td>&lt;.001 1.28 &lt;.001</td>
<td>0.064 5.22 &lt;.001</td>
<td>0.370 5.22 &lt;.001</td>
</tr>
<tr>
<td>School size $\times$ $m$</td>
<td>&lt;.001 0.77 &lt;.001</td>
<td>0.108 0.98 &lt;.001</td>
<td>0.069 0.98 &lt;.01</td>
</tr>
<tr>
<td>SSE</td>
<td>11.001</td>
<td>3,381.530</td>
<td>3,401.500</td>
</tr>
</tbody>
</table>

Note. Practically significant $\eta^2$ values ($\eta^2 > .01$) appear in bold face. $m =$ number of schools that mobile student attended. There were 47,980 error degrees of freedom. SSE = sum of squared errors. $p < .01$. 

Ignoring multiple membership.
substantially associated with this overestimation: mobility and number of schools that mobile student attended. None of the two-way interaction effects appeared to have a substantial effect.

The percentage of mobile students (see Tables 1 and 2) was a significant factor affecting the relative bias found in the level-one variance component estimates, $F(1, 47,980) = 113,807.00, p < .001, \eta^2 = .676$. The average relative bias in the estimation of level-one variance component in conditions with 10% mobile students was 0.933, as compared with 1.750 for conditions with 20% mobile students. The higher percentage of mobile students led to a substantial increase in the positive relative bias found in the level-one variance estimates. However, even with only 10% mobility, there was an unacceptably large degree of positive bias.

The number of schools that mobile students attended was also somewhat strongly associated with the relative bias, $F(1, 47,980) = 6,104.33, p < .001, \eta^2 = .036$. The average relative bias in conditions where mobile students attended two different schools was 1.247, compared to 1.436 in conditions where mobile students attended three different schools. It appears that the more schools mobile students attend, the greater the positive relative bias in the level-one variance estimate.

Use of the MM resulted in an underestimation of the school-level variability, resulting in average relative bias of $-0.1642$ (with $SD = 0.2760$). No substantial bias was found in the estimation of the level-two variance component with the use of the MMMM ($M = 0.0273, SD = 0.2323$).

The ANOVA results for MM revealed that the relative bias of the school-level variance estimates was significantly associated with the percentage of mobile students, $F(1, 47,980) = 2,555.15, p < .01, \eta^2 = .050$. The conditions in which 10% of student changed schools had an average negative bias of $-0.103$, while conditions in which 20% of students transferred schools had an average negative bias of $-0.226$ (see Tables 1 and 2). The higher the percentage of mobile students in a data set, the more negative relative bias was found. No other main effects or two-way interactions were found to have a substantial impact on the relative bias of the MM estimates.

### 3.3. Fit index

A difference of 5 or more (as suggested by Li, Bolt, & Fu, 2006) was used as a minimum cut-off that represents a substantial drop in DIC to support the model with the smaller DIC value. Using this criterion led to 100% correct model identification. DIC values estimated for the MMMM were consistently lower than the DICs estimated for the MM by 5 or more points across all 48 conditions.

### 4. Discussion

The current study compared parameter estimates and correct model identification rates of the DIC criterion when a multiple membership data structure was handled using the MMMM and using an MM that ignored the multiple membership structure. Some important differences were found and will be discussed here along with their implications.

Overall, the results of the simulation study indicated that appropriate modelling of multiple membership data structures through the use of the MMMM did not lead to biased estimates across the set of conditions examined. However, use of a conventional MM that recognized only the last school attended led to bias in the estimation of the
Ignoring multiple membership

There is a simple explanation for some of the bias noted in the level-two predictor coefficient estimates. Values on the level-two predictor, $Z$, for each level-one unit $i$ were generated using equation (2) as a weighted sum of the values on $Z$ for each level-two unit in the set $\{j\}$ associated with level-one unit $i$. This was designed to mimic a situation in which there truly is an effect of a school-level descriptor on a student outcome. However, when the MM was estimated, the value used for $Z$ (see equation (4)) reflected only the value on the school descriptor for the last school attended. For mobile students, the value used for $Z$ when the MM was estimated did not match the true value on $Z$. This clearly would have a negative impact on the resulting coefficient, thereby reducing the strength of the relationship between $Z$ and the outcome.

Use of the MMMM resulted in no substantial estimation bias detected in either the level-one or level-two variance components. However, severe overestimation of the level-one variance component was identified here when the MM was assumed. The two factors contributing to the extent of the multiple membership (here, of student mobility) seemed strongly related to this bias; the higher the proportion of mobile students and the larger the number of schools that mobile students attended, the more positively biased were the level-one variance component estimates. It seems likely that some of this bias is attributable to the use of the incorrect $Z$ values with the MM. However, use of the MM led to underestimation of the level-two variance component. Thus, the variability that truly had to do with level-two units seems to have been apportioned instead to variability within level-one units. This reapportioning of variance matches the pattern identified in related cross-classified multilevel model (CCMM) research (e.g., Luo & Kwok, 2009; Meyers & Beretvas, 2006).

In terms of the underestimation of the level-two variance component by the MM, recall that when using the MMMM, the residuals for the set of level-two units ($u_{0j}$) are assumed normally distributed with a constant variance, $\tau_{00}$. Given that the true school-level variance ($\sigma^2$) was generated to fit the MMMM, the actual contribution of the set of schools $\{j\}$ attended by a mobile student to the school-level variance is assumed to be

$$\text{var} \left( \sum_{b \in \{j\}} w_{ib} u_{0b} \right) = \tau_{00} \sum_{b \in \{j\}} (w_{ib})^2. \quad (6)$$

Given that each level-one unit’s weights sum to 1, the variance of the resulting weighted sum of level-two residuals will always be smaller than the variance of the level-two residuals themselves (Goldstein, 2010; Leckie, 2009). The MMMM takes into account the relative contributions of the school effects to the student outcome. However, the MM does not reflect the reduced contribution of mobile students’ schools to the school-level variance component, resulting in underestimation of the school-level variance component. In a data set in which, for example, all students are mobile attending two schools with weights of .5, the underestimation will be of the order of 50%. In the data sets simulated in the current study and designed to provide an authentic match to the degree of mobility typically encountered in educational settings, the degree of underestimation bias will be diluted by the large degree of non-mobile students. This helps explain the relationship between the degree of mobility and the degree of underestimation bias identified for the level-two variance component. This
result also matches what other researchers (e.g., Goldstein, 2010; Leckie, 2009) have noted.

4.1. Limitations and suggestions for future research

This study was a preliminary investigation designed to assess how parameter estimates from the MMMM designed to handle multiple membership data structures differ from estimates from use of an MM that ignores multiple membership. Given that this is the first study quantifying the potential biases, there were several limitations.

Although the parameter values used in the simulation study were selected using values encountered in both applied and methodological research, the findings presented in the simulation study should not be generalized before investigating more broadly how these factors affect MMMM estimation. The current simulation study generated data to fit a particularly simple MMMM with only one (fixed) level-one predictor and one level-two predictor. Use of the MMMM under the conditions examined resulted in unbiased estimates of this small set of parameters. Future research could investigate parameter recovery for more complex MMMMs including scenarios with misspecified MMMMs.

More importantly, it should be emphasized that previous research (Browne & Draper, 2006) has indicated that use of MCMC estimation can lead to positively biased estimates of the higher-level variance components when there are only a small number of higher-level units. The current study found that when a data set included 50 or 100 level-two units, use of MCMC estimation did not lead to biased level-two variance component estimates. However, we also ran some additional simulations using 30 level-two units and encountered substantial level-two variance component bias with the MMMM estimates. Thus, applied researchers should be careful to ensure that they have a sufficiently large number of level-two units when using MCMC estimation to estimate the MMMM.

Another limitation of the current study is that we examined a simple model including a fixed student-level predictor whose effect was modelled (both in the generating and estimating models) as not varying across schools. In addition, the students’ actual scores on the student predictor were not generated to differ across schools. However, in real data, nearly all student characteristics are likely to vary somewhat among schools as students are not randomly assigned to schools. For example, for a variety of reasons, some schools have higher proportions of students of high socio-economic status or greater ability than do others. In addition, the effect of a student characteristic (e.g., a pre-test score) on their later achievement might vary by school. For example, some schools might have better instructional programmes that lessen the influence of poor prior achievement on later achievement test scores. Thus, although the current study found that the student predictor coefficient was not substantially biased, this finding should not be generalized to scenarios in which there might be variability among schools on the student characteristic of interest (e.g., socio-economic status or pre-test scores) or in which the influence of the student characteristic on the outcome (e.g., achievement) might vary by school. It seems possible that the coefficient for the level-one predictors might then be biased in these more realistic scenarios especially when multiple membership is ignored. Future research should be conducted that assesses MMMM recovery of level-one predictor coefficients under these more realistic conditions.

The current study should also be extended to include a third level that would permit the modelling of student growth trajectories. For example, in growth modelling, time points (level one) are nested within students (level two) and students might be members of multiple schools (level three). This extension would be useful given the current
educational interest in tracking students’ yearly progress. In addition, residuals are increasingly being used to make important decisions about the value added by schools and classrooms, and while some work has been done on this issue (see Leckie, 2009), more research is needed to assess the effect of using the MMMM and the MM on residuals estimated using data sets that include mobile students.

Finally, note that the current simulation study was run using MLwiN software to generate data and estimate model parameters; SAS code was used to summarize the results. A module developed by Leckie and Charlton (2011) was recently made available which allows MLwiN to be run completely within Stata software without having to write MLwiN macros. Use of this runmlwin module might facilitate future empirical simulation studies designed to extend this current line of research.

4.2. Implications
The incorrect shift in variability from level-two to level-one units when the MM was used with multiple membership data sets will lead to substantially underestimated ICC values. These low, underestimated ICC values could mislead researchers to select traditional regression models over multilevel models, resulting in further inferential errors typically associated with this misspecification (see Raudenbush & Bryk, 2002). These negatively biased ICC values could also have important implications for decisions being made about schools. Underestimated ICC values might mislead researchers to underemphasize the contribution and potentially thus the importance of the level-two unit (e.g., school) to the level-one outcome (e.g., student). Along the same lines, failure to model prior schools attended also resulted in underestimation of the school-level predictor’s coefficient. Underestimation of school-level predictors’ effects could also lead to misinterpretation of schools’ effects on academic achievement.

The findings of the current research study underscore what is already known about the effects of ignoring multiple membership structures (Goldstein, 2010; Leckie, 2009) and were intended to quantify some of this negative impact by simulating realistic degrees of mobility. The underestimation of the level-two variance component and of the level-two predictor coefficient when multiple membership structures are ignored constitutes one of the primary findings of the study due to its particular relevance in educational research. Schools are being held accountable for students’ performance and thus optimally performing estimation of school effects is crucial (Leckie & Goldstein, 2009).

There are alternative approaches to the use of the MMMM for handling mobile students’ data. For example, the CCMM can be used with each classification representing a single school attended for each time period of interest (e.g., school year). The resulting CCMM corresponds with the modelling approach on which the Tennessee Value-Added Assessment System (TVAAS) is based. However, use of the CCMM is limited to scenarios in which there is no mobility within each time period of interest, although there is mobility across time periods (Beretvas & Leite, 2010). In addition, the assumption when using the CCMM in this way would be that the timing of any school moves corresponds exactly with the classification factors’ set-up. In other words, use of the CCMM would require that there be as many classifications as the maximum number of schools attended by a mobile student in the data set. Unfortunately, while more sophisticated models used in large-scale assessments recognize yearly school changes, the models and associated analyses do not typically handle within-year mobility, thus leading to the same kind of inferential errors as were noted in the current study.
In light of the implications of findings of the current study for decisions being made about schools and their effects on student achievement, it behoves educational researchers and practitioners to better understand use of the MMMM with multiple membership data structures. It is hoped that research on and use of approaches for handling multiple membership structures will continue and will lead to more accurate inferences being made about the relevant constituents.

References


Received 18 August 2010; revised version received 19 May 2011
The correspondence between Beretvas’s (2010) multiple membership multilevel model (MMMM) formulation and that of Browne, Goldstein and Rasbash (2001) is as shown in the following table.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level-one equation</strong></td>
<td>$Y_{i(j)} = \beta_0(j) + \beta_1(j) X_{i(j)} + e_{i(j)}$</td>
<td>$Y_i = \beta_0 + \beta_2 X_i + e_i$</td>
</tr>
<tr>
<td></td>
<td>$\beta_0(j) = \gamma_{00} + \gamma_{10} \sum_{b \in {j}} w_{ib} Z_b + \sum_{b \in {j}} w_{ib} U_{0b}$</td>
<td>$\beta_0 = \beta_0 + \beta_1 \sum_{j \in \text{Sch}(i)} w_{i(j)} Z_{j(2)} + \sum_{j \in \text{Sch}(i)} w_{i(j)} U_{j(2)}$</td>
</tr>
<tr>
<td></td>
<td>$\beta_1(j) = \gamma_{01}$</td>
<td>$\beta_2 = \beta_2$</td>
</tr>
<tr>
<td><strong>Level-two equation</strong></td>
<td>$Y_{i(j)} = \gamma_{00} + \gamma_{10} \sum_{b \in {j}} w_{ib} Z_b + \gamma_{11} X_{i(j)} + \sum_{b \in {j}} w_{ib} U_{0b} + e_{i(j)}$</td>
<td>$Y_i = \beta_0 + \beta_1 \sum_{j \in \text{Sch}(i)} w_{i(j)} Z_{j(2)} + \beta_2 X_i + \sum_{j \in \text{Sch}(i)} w_{i(j)} U_{j(2)} + e_i$</td>
</tr>
<tr>
<td><strong>Combined equation</strong></td>
<td>$Y_{i(j)} = \gamma_{00} + \gamma_{10} \sum_{b \in {j}} w_{ib} Z_b + \gamma_{11} X_{i(j)} + \sum_{b \in {j}} w_{ib} U_{0b} + e_{i(j)}$</td>
<td>$Y_i = \beta_0 + \beta_1 \sum_{j \in \text{Sch}(i)} w_{i(j)} Z_{j(2)} + \beta_2 X_i + \sum_{j \in \text{Sch}(i)} w_{i(j)} U_{j(2)} + e_i$</td>
</tr>
<tr>
<td><strong>Assumptions</strong></td>
<td>$u_{0b} \sim N (0, \tau_{00})$</td>
<td>$u_{j(2)} \sim N (0, \sigma_{U_{j(2)}})$</td>
</tr>
<tr>
<td></td>
<td>$e_{i(j)} \sim N (0, \sigma^2)$</td>
<td>$e_i \sim N (0, \sigma^2)$</td>
</tr>
</tbody>
</table>

*Note. Sch = School.*
学霸图书馆
www.xuebalib.com

本文献由“学霸图书馆-文献云下载”收集自网络，仅供学习交流使用。

学霸图书馆（www.xuebalib.com）是一个“整合众多图书馆数据库资源，提供一站式文献检索和下载服务”的24小时在线不限IP图书馆。

图书馆致力于便利、促进学习与科研，提供最强文献下载服务。

图书馆导航：
图书馆首页    文献云下载    图书馆入口    外文数据库大全    疑难文献辅助工具